

(i) APPLICANT: Ciossek, Thomas
Ullrich, Axel
Millauer, Birgit

ii) TITLE OF INVENTION: METHODS FOR DIAGNOSIS
AND TREATMENT OF MDK1
SIGNAL TRANSDUCTION
DISORDERS

iii) NUMBER OF SEQUENCES: 12

iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Lyon & Lyon
(B) STREET: 633 West Fifth Street
Suite 4700
(C) CITY: Los Angeles
(D) STATE: California
(E) COUNTRY: U.S.A.
(F) ZIP: 90071-2066

(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

(B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: IBM P.C. DOS 5.0

(D) SOFTWARE: Word Perfect 5.1

(A) APPLICATION NUMBER: 08/368,776
(B) FILING DATE: January 3, 1995

(vii) PRIOR APPLICATION DATA:

Prior applications total,
including application
described below: none

- (A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Warburg, Richard J.
(B) REGISTRATION NUMBER: 32,327
(C) REFERENCE/DOCKET NUMBER: 208/007

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (213) 489-1600
(B) TELEFAX: (213) 955-0440
(C) TELEX: 67-3510

(2) INFORMATION FOR SEQUENCE ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4304 base pairs
(B) TYPE: nucleic acid
(C) STRANDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

AAGCGGCCGG TCTGCAGTCG GAGACTTGCA GGCAGCAAAC ACGGTGCGAA      50
CGAACCCGAG GGGGGAGAGA GAAATCAAAC AGCTAAGCGT GGAGCAGACG      100
GCCTGGGACC CAGAAGGGGA TCGATGCGAG GAGCGCAATA ATAACAACAA      150
TAATAACCCA CTTCGGAGCA AACAGCATCT AAAGAGCTGC GACCCAACTG      200
CAGCCTAAAA AAATCAAACC TGCTCATGCA CCATGGTTGT TCAAACTCGG      250
TTCCCTTCGT GGATTATTTT GTGTTACATC TGGCTGCTTG GCTTTGCACA      300
CACGGGGGAG GCGCAGGCTG CGAAGGAAGT ACTATTACTG GACTCGAAAG      350
CACAACAAAC AGAATTGGAA TGGATTTTCT CTCCACCCAG TGGGTGGGAA      400
GAAATTAGTG GTTTGGATGA GAACTACACT CCGATAAGAA CATACCAGGT      450
GTGCCAGGTC ATGGAGCCCA ACCAGAACAA CTGGCTGCGG ACTAACTGGA      500
TTTCTAAAGG CAACGCACAA AGGATTTTGT TAGAATTGAA ATTCACCTTG      550
AGGGATTGTA ATAGTCTTCC CGGAGTCCTG GGAAGTTGCA AGGAAACGTT      600

```

TAATTTGTAC TATTATGAAA CAGACTACGA CACCGGCAGG AATATACGAG	650
AAAACCTTTA TGTTAAAATA GACACCATTG CTGCAGATGA AAGTTTCACA	700
CAAGGTGACC TTGGTGAAAG AAAGATGAAG CTGAACACTG AGGTGAGAGA	750
GATTGGACCT TTGTCCAAAA AGGGATTCTA TCTTGCCTTT CAGGATGTAG	800
GGGCTTGCAAT AGCATTGGTT TCTGTCAAAG TGTACTACAA GAAGTGCTGG	850
ACCATTGTTG AGAACTTAGC TGTCTTTCCA GATACAGTGA CTGGTTCGGA	900
ATTTTCCTCC TTAGTCGAGG TCCGTGGGAC ATGTGTCAGC AGTGCCGAGG	950
AAGAGGCAGA AAATTCCCCC AGAATGCATT GCAGTGCAGA AGGAGAGTGG	1000
CTAGTACCCA TTGGAAAATG CATCTGCAAA GCAGGCTATC AGCAAAAAGG	1050
GGACACTTGC GAACCCTGTG GCCGCAGGTT CTACAAATCT TCCTCTCAGG	1100
ATCTCCAGTG TTCTCGTTGT CCAACCCACA GCTTCTCTGA CCGAGAAGGA	1150
TCATCCAGGT GTGAATGTGA AGATGGGTAC TACAGAGCTC CTTCTGATCC	1200
ACCATACGTT GCATGCACGA GGCCTCCCTC TGCACCACAG AACCTTATTT	1250
TCAATATCAA TCAAACGACT GTAAGTTTGG AATGGAGTCC TCCGGCTGAC	1300
AACGGGGGAA GAAACGATGT CACCTACAGA ATACTGTGTA AGCGGTGCAG	1350
TTGGGAACAG GGAGAATGTG TGCCATGCGG AAGTAACATT GGATACATGC	1400
CCCAGCAGAC GGGATTAGAG GATAACTATG TCACTGTCAT GGACCTACTT	1450
GCCCATGCAA ATTACACTTT CGAAGTTGAA GCTGTAAATG GAGTTTCGGA	1500
CTTAAGCAGA TCCCAGAGGC TCTTCGCTGC TGTTAGCATC ACCACCGGTC	1550
AAGCAGCTCC CTCGCAAGTG AGTGGAGTCA TGAAGGAGCG AGTACTGCAG	1600
CGGAGTGTGC AGCTTTCCTG GCAGGAGCCG GAGCATCCCA ATGGAGTCAT	1650
CACGGAATAT GAAATCAAGT ATTATGAGAA AGATCAACGG GAAAGGACGT	1700
ACTCAACACT CAAAACCAAG TCCACCTCCG CCTCCATTAA TAATCTGAAA	1750
CCGGGAACAG TGTACGTCTT TCAGATCCGG GCGGTCACTG CTGCCGGTTA	1800
TGGAAACTAC AGCCCTAGGC TTGATGTTGC CACACTTGAG GAAGCTTCAG	1850
GTAAAATGTT TGAAGCGACA GCAGTCTCCA GTGAACAGAA TCCTGTCATC	1900
ATAATTGCTG TAGTGGCTGT AGCAGGGACC ATCATCTTGG TGTTCATGGT	1950
GTTCCGGCTTC ATCATTGGAA GAAGGCACTG TGGTTATAGC AAGGCTGACC	2000

AAGAAGGGGA TGAAGAACTC TACTTTCATT TTAAATTTCC AGGCACCAAA	2050
ACCTACATTG ACCCTGAAAC CTATGAGGAC CCAAATAGAG CTGTCCATCA	2100
ATTCGCCAAG GAGCTAGATG CCTCCTGTAT TAAAATTGAG CGTGTGATTG	2150
GTGCAGGAGA ATTTGGAGAA GTTTCAGTG GTCGTTTGA ACTTCCGGGC	2200
CAGAGAGATG TTGCAGTGGC CATAAAAACC CTGAAAGTTG GTTACACAGA	2250
AAAGCAAAGG AGGGACTTTT TATGCGAAGC AAGCATCATG GGGCAATTTG	2300
ACCACCCAAA TGTCGTCCAT TTGGAAGGGG TTGTTACAAG AGGGAAGCCT	2350
GTCATGATTG TGATAGAGTT CATGGAGAAT GGAGCCCTGG ATGCATTTCT	2400
CAGGAAACAC GATGGGCAGT TTACAGTCAT TCAGTTGGTA GGAATGTTGA	2450
GAGGTATTGC CGCTGGGATG CGATACTTGG CTGATATGGG ATACGTTTAC	2500
AGGGACCTTG CAGCGCGCAA CATCCTTGTC AACAGCAATC TTGTTTGTAA	2550
AGTGTGAGAT TTTGGCCTTT CCCGGGTTAT AGAGGATGAT CCCGAAGCTG	2600
TCTACACCAC GACTGGTGGA AAAATTCCAG TAAGGTGGAC TGCACCGGAA	2650
GCCATTCAAT ACCGGAAGTT CACCTCAGCC AGCGATGTGT GGAGCTATGG	2700
GATTGTGATG TGGGAAGTGA TGTCTTATGG AGAAAGACCT TACTGGGACA	2750
TGTCAAATCA AGATGTCATT AAAGCGATAG AAGAAGGTTA TCGTTTGCCG	2800
GCGCCCATGG ATGCCCAGC TGGTCTTCAC CAGCTAATGC TGGATTGTTG	2850
GCAGAAAGAT CGGGCGGAAA GGCCAAAGTT TGAGCAGATA GTCGGAATTC	2900
TAGACAAAAT GATTCGAAAC CCAAGTAGTC TGAAAACACC CCTGGGAACT	2950
TGTAGTAGAC CCTTAAGCCC TCTTCTGGAC CAGAGCACTC CTGACTTCAC	3000
TGCCTTCTGT TCAGTTGGAG AATGGTTGCA AGCTATTAAA ATGGAAAGGT	3050
ATAAGGACAA CTTACAGCA GCGGGTTACA ACTCACTCGA GTCAGTGGCC	3100
AGGATGACTA TCGATGATGT GATGAGTTTA GGGATCACAC TGTTGGCCA	3150
TCAAAGAAG ATCATGAGCA GCATCCAGAC TATGCGGGCA CAAATGTTGC	3200
ATTTACACGG AACAGGCATC CAAGTGTGAC ACATCGGCCT CCCTCAGATG	3250
AGGCTTAAGA CTGCAGGAGA ACAGTTCCTGG CCTTCAGTAT ACGCATAGAA	3300
TGCTGCTAGA AGACAGTTGA TATACTGGGT CCTTCCTACA AGAAAGAGAA	3350
GATTTTAGAA GCACCTCCAG ACTTGAAGTC CTAAGTGCCA CCAGAATATA	3400

CAAAAAGGGA ATTTAGGATC CACCACTGGT GGCCAGGAAC ACAGCAGAGA	3450
CAATAAACAA AGTACTACCT GAAAAACATC CCAACACCTT GAGCTCTCGA	3500
ACCTCCTTTT TATCTTATAG ACTTTTAAAA AATGTACATA AAGAATTTAA	3550
GAAAGAATAT ATTTGTCAAA TAAAAATCAT GATCTTATTG TTAAAATCAA	3600
TGAAATATTT TCCTTAAAAAT ATGTGATTTT AGACTATTCT TTTCCAGAAC	3650
CATCTGTGTT TATTCTGCTT AAGGACTTTG TTTTAGAAAG TTATTTGTAG	3700
CTTTGGACCT TTTTAGTGTT AAATTTATGA CACGTTACTA CACTGGGAAC	3750
CTTTGAAGAC TCTCAAACCT AAAGGAAAGC AAAACTACGC ACATAGTCGA	3800
GGATGGACTT TGTCCTTCAT GGCTTTGGTA TCCTGGCTGT GTCATTTTGT	3850
TAAACCAAGT ATGTTTTTCAT ATTGTTTGCT GATTGGCAGG TAGTTCAAAA	3900
TGCAAGTTG CCAAGAGCTC TGATATTTTT TAACAGGATT TTTTTTCTT	3950
TGTAAAAATC AGATAACATA CTAACTTTTC AATGAAAAAA AAAAAAAAAG	4000
AAGCAATAAT GATCCATAAA TACTATAAGG CACTTTTAAC AGATTGTTTA	4050
TAGAGTGATT TACTAGGCAG AATTTAATAA AAAAAAAGA GAGATGTCAA	4100
ATTTTAGGTT TATGTGTATA TGATAAAAGG CTGAGCTTCG TCTGAAGATG	4150
CTGGTGAAAG CAAGACTGGA AGCGAAGCTC TCCAGCTTTG GCTAACCCAA	4200
TCCGAGCACA TCAAGAGCTT CAGTCTGTG ACAGTAAGAA ATTTAGGAAC	4250
ATAGTTGACC TATATTTTGT ATTCTTTCTT GTTGAATGCA GTCCAAATAC	4300
AAAA	4304

(2) INFORMATION FOR SEQUENCE ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH:	998 amino acids
(B) TYPE:	amino acid
(C) STRANDNESS:	single
(D) TOPOLOGY:	linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Val Val Gln Thr Arg Phe Pro Ser Trp Ile Ile Leu Cys Tyr Ile
1 5 10 15

Trp Leu Leu Gly Phe Ala His Thr Gly Glu Ala Gln Ala Ala Lys Glu
20 25 30

Val Leu Leu Leu Asp Ser Lys Ala Gln Gln Thr Glu Leu Glu Trp Ile
35 40 45

Ser Ser Pro Pro Ser Gly Trp Glu Glu Ile Ser Gly Leu Asp Glu Asn
50 55 60

Tyr Thr Pro Ile Arg Thr Tyr Gln Val Cys Gln Val Met Glu Pro Asn
65 70 75 80

Gln Asn Asn Trp Leu Arg Thr Asn Trp Ile Ser Lys Gly Asn Ala Gln
85 90 95

Arg Ile Phe Val Glu Leu Lys Phe Thr Leu Arg Asp Cys Asn Ser Leu
100 105 110

Pro Gly Val Leu Gly Thr Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr
115 120 125

Glu Thr Asp Tyr Asp Thr Gly Arg Asn Ile Arg Glu Asn Leu Tyr Val
130 135 140

Lys Ile Asp Thr Ile Ala Ala Asp Glu Ser Phe Thr Gln Gly Asp Leu
145 150 155 160

Gly Glu Arg Lys Met Lys Leu Asn Thr Glu Val Arg Glu Ile Gly Pro
165 170 175

Leu Ser Lys Lys Gly Phe Tyr Leu Ala Phe Gln Asp Val Gly Ala Cys
180 185 190

Ile Ala Leu Val Ser Val Lys Val Tyr Tyr Lys Lys Cys Trp Thr Ile
195 200 205

Val Glu Asn Leu Ala Val Phe Pro Asp Thr Val Thr Gly Ser Glu Phe
210 215 220

Ser Ser Leu Val Glu Val Arg Gly Thr Cys Val Ser Ser Ala Glu Glu
225 230 235 240

Glu Ala Glu Asn Ser Pro Arg Met His Cys Ser Ala Glu Gly Glu Trp
245 250 255

Leu Val Pro Ile Gly Lys Cys Ile Cys Lys Ala Gly Tyr Gln Gln Lys
260 265 270

Gly Asp Thr Cys Glu Pro Cys Gly Arg Arg Phe Tyr Lys Ser Ser Ser
 275 280 285
 Gln Asp Leu Gln Cys Ser Arg Cys Pro Thr His Ser Phe Ser Asp Arg
 290 295 300
 Glu Gly Ser Ser Arg Cys Glu Cys Glu Asp Gly Tyr Tyr Arg Ala Pro
 305 310 315 320
 Ser Asp Pro Pro Tyr Val Ala Cys Thr Arg Pro Pro Ser Ala Pro Gln
 325 330 335
 Asn Leu Ile Phe Asn Ile Asn Gln Thr Thr Val Ser Leu Glu Trp Ser
 340 345 350
 Pro Pro Ala Asp Asn Gly Gly Arg Asn Asp Val Thr Tyr Arg Ile Leu
 355 360 365
 Cys Lys Arg Cys Ser Trp Glu Gln Gly Glu Cys Val Pro Cys Gly Ser
 370 375 380
 Asn Ile Gly Tyr Met Pro Gln Gln Thr Gly Leu Glu Asp Asn Tyr Val
 385 390 395 400
 Thr Val Met Asp Leu Leu Ala His Ala Asn Tyr Thr Phe Glu Val Glu
 405 410 415
 Ala Val Asn Gly Val Ser Asp Leu Ser Arg Ser Gln Arg Leu Phe Ala
 420 425 430
 Ala Val Ser Ile Thr Thr Gly Gln Ala Ala Pro Ser Gln Val Ser Gly
 435 440 445
 Val Met Lys Glu Arg Val Leu Gln Arg Ser Val Gln Leu Ser Trp Gln
 450 455 460
 Glu Pro Glu His Pro Asn Gly Val Ile Thr Glu Tyr Glu Ile Lys Tyr
 465 470 475 480
 Tyr Glu Lys Asp Gln Arg Glu Arg Thr Tyr Ser Thr Leu Lys Thr Lys
 485 490 495
 Ser Thr Ser Ala Ser Ile Asn Asn Leu Lys Pro Gly Thr Val Tyr Val
 500 505 510
 Phe Gln Ile Arg Ala Val Thr Ala Ala Gly Tyr Gly Asn Tyr Ser Pro
 515 520 525
 Arg Leu Asp Val Ala Thr Leu Glu Glu Ala Ser Gly Lys Met Phe Glu
 530 535 540
 Ala Thr Ala Val Ser Ser Glu Gln Asn Pro Val Ile Ile Ile Ala Val
 545 550 555 560

110

Val Ala Val Ala Gly Thr Ile Ile Leu Val Phe Met Val Phe Gly Phe
565 570 575

Ile Ile Gly Arg Arg His Cys Gly Tyr Ser Lys Ala Asp Gln Glu Gly
580 585 590

Asp Glu Glu Leu Tyr Phe His Phe Lys Phe Pro Gly Thr Lys Thr Tyr
595 600 605

Ile Asp Pro Glu Thr Tyr Glu Asp Pro Asn Arg Ala Val His Gln Phe
610 615 620

Ala Lys Glu Leu Asp Ala Ser Cys Ile Lys Ile Glu Arg Val Ile Gly
625 630 635 640

Ala Gly Glu Phe Gly Glu Val Cys Ser Gly Arg Leu Lys Leu Pro Gly
645 650 655

Gln Arg Asp Val Ala Val Ala Ile Lys Thr Leu Lys Val Gly Tyr Thr
660 665 670

Glu Lys Gln Arg Arg Asp Phe Leu Cys Glu Ala Ser Ile Met Gly Gln
675 680 685

Phe Asp His Pro Asn Val Val His Leu Glu Gly Val Val Thr Arg Gly
690 695 700

Lys Pro Val Met Ile Val Ile Glu Phe Met Glu Asn Gly Ala Leu Asp
705 710 715 720

Ala Phe Leu Arg Lys His Asp Gly Gln Phe Thr Val Ile Gln Leu Val
725 730 735

Gly Met Leu Arg Gly Ile Ala Ala Gly Met Arg Tyr Leu Ala Asp Met
740 745 750

Gly Tyr Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser
755 760 765

Asn Leu Val Cys Lys Val Ser Asp Phe Gly Leu Ser Arg Val Ile Glu
770 775 780

Asp Asp Pro Glu Ala Val Tyr Thr Thr Thr Gly Gly Lys Ile Pro Val
785 790 795 800

Arg Trp Thr Ala Pro Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala
805 810 815

Ser Asp Val Trp Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr
820 825 830

Gly Glu Arg Pro Tyr Trp Asp Met Ser Asn Gln Asp Val Ile Lys Ala
835 840 845

SSSD/27372. v01

111

Ile Glu Glu Gly Tyr Arg Leu Pro Ala Pro Met Asp Cys Pro Ala Gly
850 855 860
Leu His Gln Leu Met Leu Asp Cys Trp Gln Lys Asp Arg Ala Glu Arg
865 870 875 880
Pro Lys Phe Glu Gln Ile Val Gly Ile Leu Asp Lys Met Ile Arg Asn
885 890 895
Pro Ser Ser Leu Lys Thr Pro Leu Gly Thr Cys Ser Arg Pro Leu Ser
900 905 910
Pro Leu Leu Asp Gln Ser Thr Pro Asp Phe Thr Ala Phe Cys Ser Val
915 920 925
Gly Glu Trp Leu Gln Ala Ile Lys Met Glu Arg Tyr Lys Asp Asn Phe
930 935 940
Thr Ala Ala Gly Tyr Asn Ser Leu Glu Ser Val Ala Arg Met Thr Ile
945 950 955 960
Asp Asp Val Met Ser Leu Gly Ile Thr Leu Val Gly His Gln Lys Lys
965 970 975
Ile Met Ser Ser Ile Gln Thr Met Arg Ala Gln Met Leu His Leu His
980 985 990
Gly Thr Gly Ile Gln Val
995

(2) INFORMATION FOR SEQUENCE ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 610 amino acids
(B) TYPE: amino acid
(C) STRANDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Val Val Gln Thr Arg Phe Pro Ser Trp Ile Ile Leu Cys Tyr Ile
1 5 10 15
Trp Leu Leu Gly Phe Ala His Thr Gly Glu Ala Gln Ala Ala Lys Glu
20 25 30
Val Leu Leu Leu Asp Ser Lys Ala Gln Gln Thr Glu Leu Glu Trp Ile
35 40 45

SSSD/27372. v01

112

Ser 50	Pro	Pro	Ser	Gly 55	Trp	Glu	Glu	Ile	Ser	Gly 60	Leu	Asp	Glu	Asn
Tyr 65	Thr	Pro	Ile	Arg 70	Thr	Tyr	Gln	Val	Cys 75	Gln	Val	Met	Glu	Pro 80
Gln	Asn	Asn	Trp	Leu 85	Arg	Thr	Asn	Trp	Ile 90	Ser	Lys	Gly	Asn	Ala 95
Arg	Ile	Phe 100	Val	Glu	Leu	Lys	Phe	Thr 105	Leu	Arg	Asp	Cys 110	Asn	Ser 115
Pro	Gly 115	Val	Leu	Gly	Thr	Cys	Lys 120	Glu	Thr	Phe	Asn	Leu 125	Tyr	Tyr 130
Glu 130	Thr	Asp	Tyr	Asp	Thr	Gly 135	Arg	Asn	Ile	Arg	Glu 140	Asn	Leu	Tyr 145
Lys 145	Ile	Asp	Thr	Ile	Ala 150	Ala	Asp	Glu	Ser	Phe 155	Thr	Gln	Gly	Asp 160
Gly	Glu	Arg	Lys	Met 165	Lys	Leu	Asn	Thr	Glu 170	Val	Arg	Glu	Ile	Gly 175
Leu	Ser	Lys	Lys 180	Gly	Phe	Tyr	Leu	Ala 185	Phe	Gln	Asp	Val	Gly 190	Ala 195
Ile	Ala 195	Leu	Val	Ser	Val	Lys	Val 200	Tyr	Tyr	Lys	Lys	Cys 205	Trp	Thr 210
Val 210	Glu	Asn	Leu	Ala	Val	Phe 215	Pro	Asp	Thr	Val	Thr 220	Gly	Ser	Glu 225
Ser 225	Ser	Leu	Val	Glu	Val 230	Arg	Gly	Thr	Cys 235	Val	Ser	Ser	Ala	Glu 240
Glu	Ala	Glu	Asn	Ser 245	Pro	Arg	Met	His	Cys 250	Ser	Ala	Glu	Gly	Glu 255
Leu	Val	Pro	Ile 260	Gly	Lys	Cys	Ile	Cys 265	Lys	Ala	Gly	Tyr	Gln 270	Gln 275
Gly	Asp	Thr 275	Cys	Glu	Pro	Cys	Gly 280	Arg	Arg	Phe	Tyr	Lys 285	Ser	Ser 290
Gln 290	Asp	Leu	Gln	Cys	Ser	Arg 295	Cys	Pro	Thr	His	Ser 300	Phe	Ser	Asp 305
Glu 305	Gly	Ser	Ser	Arg	Cys 310	Glu	Cys	Glu	Asp	Gly 315	Tyr	Tyr	Arg	Ala 320
Ser	Asp	Pro	Pro	Tyr 325	Val	Ala	Cys	Thr	Arg 330	Pro	Pro	Ser	Ala	Pro 335

SSSD/27372. v01

113

Asn Leu Ile Phe Asn Ile Asn Gln Thr Thr Val Ser Leu Glu Trp Ser
340 345 350

Pro Pro Ala Asp Asn Gly Gly Arg Asn Asp Val Thr Tyr Arg Ile Leu
355 360 365

Cys Lys Arg Cys Ser Trp Glu Gln Gly Glu Cys Val Pro Cys Gly Ser
370 375 380

Asn Ile Gly Tyr Met Pro Gln Gln Thr Gly Leu Glu Asp Asn Tyr Val
385 390 395 400

Thr Val Met Asp Leu Leu Ala His Ala Asn Tyr Thr Phe Glu Val Glu
405 410 415

Ala Val Asn Gly Val Ser Asp Leu Ser Arg Ser Gln Arg Leu Phe Ala
420 425 430

Ala Val Ser Ile Thr Thr Gly Gln Ala Ala Pro Ser Gln Val Ser Gly
435 440 445

Val Met Lys Glu Arg Val Leu Gln Arg Ser Val Gln Leu Ser Trp Gln
450 455 460

Glu Pro Glu His Pro Asn Gly Val Ile Thr Glu Tyr Glu Ile Lys Tyr
465 470 475 480

Tyr Glu Lys Asp Gln Arg Glu Arg Thr Tyr Ser Thr Leu Lys Thr Lys
485 490 495

Ser Thr Ser Ala Ser Ile Asn Asn Leu Lys Pro Gly Thr Val Tyr Val
500 505 510

Phe Gln Ile Arg Ala Val Thr Ala Ala Gly Tyr Gly Asn Tyr Ser Pro
515 520 525

Arg Leu Asp Val Ala Thr Leu Glu Glu Ala Ser Gly Lys Met Phe Glu
530 535 540

Ala Thr Ala Val Ser Ser Glu Gln Asn Pro Val Ile Ile Ile Ala Val
545 550 555 560

Val Ala Val Ala Gly Thr Ile Ile Leu Val Phe Met Val Phe Gly Phe
565 570 575

Ile Ile Gly Arg Arg His Cys Gly Tyr Ser Lys Ala Asp Gln Glu Gly
580 585 590

Asp Glu Glu Leu Tyr Phe His Ser Leu Val Thr Asn Glu His Leu Ser
595 600 605

Val Leu
610

SSSD/27372. v01

(2) INFORMATION FOR SEQUENCE ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2901 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

AAGCGGCCGG TCTGCAGTCG GAGACTTGCA GGCAGCAAAC ACGGTGCGAA	50
CGAACCGGAG GGGGGAGAGA GAAATCAAAC AGCTAAGCGT GGAGCAGACG	100
GCCTGGGACC CAGAAGGGGA TCGATGCGAG GAGCGCAATA ATAACAACAA	150
TAATAACCCA CTTCGGAGCA AACAGCATCT AAAGAGCTGC GACCCAACTG	200
CAGCCTAAAA AAATCAAACC TGCTCATGCA CCATGGTTGT TCAAACTCGG	250
TTCCCTTCGT GGATTATTTT GTGTTACATC TGGCTGCTTG GCTTTGCACA	300
CACGGGGGAG GCGCAGGCTG CGAAGGAAGT ACTATTACTG GACTCGAAAG	350
CACAACAAAC AGAATTGGAA TGGATTTCTT CTCCACCCAG TGGGTGGGAA	400
GAAATTAGTG GTTTGGATGA GAACTACACT CCGATAAGAA CATAACAGGT	450
GTGCCAGGTC ATGGAGCCCA ACCAGAACAA CTGGCTGCGG ACTAACTGGA	500
TTTCTAAAGG CAACGCACAA AGGATTTTTG TAGAATTGAA ATTCACCTTG	550
AGGGATTGTA ATAGTCTTCC CGGAGTCCTG GGAACCTGCA AGGAAACGTT	600
TAATTTGTAC TATTATGAAA CAGACTACGA CACCGGCAGG AATATACGAG	650
AAAACCTTTA TGTTAAATA GACACCATTG CTGCAGATGA AAGTTTCACA	700
CAAGGTGACC TTGGTGAAAG AAAGATGAAG CTGAACACTG AGGTGAGAGA	750
GATTGGACCT TTGTCCAAAA AGGGATTCTA TCTTGCCTTT CAGGATGTAG	800
GGGCTTGCAAT AGCATTGGTT TCTGTCAAAG TGTACTACAA GAAGTGCTGG	850
ACCATTGTTG AGAACTTAGC TGTCTTTCCA GATACAGTGA CTGGTTCGGA	900
ATTTTCCTCC TTAGTCGAGG TCCGTGGGAC ATGTGTCAGC AGTGCCGAGG	950
AAGAGGCAGA AAATTCCCCC AGAATGCATT GCAGTGCAGA AGGAGAGTGG	1000
CTAGTACCCA TTGGAAAATG CATCTGCAAA GCAGGCTATC AGCAAAAAGG	1050

GGACACTTGC	GAACCCTGTG	GCCGCAGGTT	CTACAAATCT	TCCTCTCAGG	1100
ATCTCCAGTG	TTCTCGTTGT	CCAACCCACA	GCTTCTCTGA	CCGAGAAGGA	1150
TCATCCAGGT	GTGAATGTGA	AGATGGGTAC	TACAGAGCTC	CTTCTGATCC	1200
ACCATACGTT	GCATGCACGA	GGCCTCCCTC	TGCACCACAG	AACCTTATTT	1250
TCAATATCAA	TCAAACGACT	GTAAGTTTGG	AATGGAGTCC	TCCGGCTGAC	1300
AACGGGGGAA	GAAACGATGT	CACCTACAGA	ATACTGTGTA	AGCGGTGCAG	1350
TTGGGAACAG	GGAGAATGTG	TGCCATGCGG	AAGTAACATT	GGATACATGC	1400
CCCAGCAGAC	GGGATTAGAG	GATAACTATG	TCACTGTCAT	GGACCTACTT	1450
GCCCATGCAA	ATTACACTTT	CGAAGTTGAA	GCTGTAAATG	GAGTTTCGGA	1500
CTTAAGCAGA	TCCCAGAGGC	TCTTCGCTGC	TGTTAGCATC	ACCACCGGTC	1550
AAGCAGCTCC	CTCGCAAGTG	AGTGGAGTCA	TGAAGGAGCG	AGTACTGCAG	1600
CGGAGTGTGC	AGCTTTCCTG	GCAGGAGCCG	GAGCATCCCA	ATGGAGTCAT	1650
CACGGAATAT	GAAATCAAGT	ATTATGAGAA	AGATCAACGG	GAAAGGACGT	1700
ACTCAACACT	CAAAACCAAG	TCCACCTCCG	CCTCCATTAA	TAATCTGAAA	1750
CCGGGAACAG	TGTACGTCTT	TCAGATCCGG	GCGGTCACTG	CTGCCGGTTA	1800
TGGAAACTAC	AGCCCTAGGC	TTGATGTTGC	CACACTTGAG	GAAGCTTCAG	1850
GTAAATGTT	TGAAGCGACA	GCAGTCTCCA	GTGAACAGAA	TCCTGTCATC	1900
ATAATTGCTG	TAGTGGCTGT	AGCAGGGACC	ATCATCTTGG	TGTTTCATGGT	1950
GTTCCGGCTTC	ATCATTGGAA	GAAGGCACTG	TGGTTATAGC	AAGGCTGACC	2000
AAGAAGGGGA	TGAAGAACTC	TACTTTCATT	CTTTAGTAAC	AAATGAGCAC	2050
CTGTCAGTTT	TATAAACCGC	AACAATAACT	GTTTAAGACA	ATCAATTTTG	2100
GATAAACAAAT	CAACTACAGC	AGAATAAATC	AAGATTTTAA	AGTCCCATTT	2150
TCCTTTATAC	ATTCTGCTTA	TTTTGTGTGT	ATATGTTTAT	TTTTTAAACT	2200
CTGATCTTGA	TTGAATGTGA	TACCATAAGC	ACAGTTAGGC	TGCAGTGTA	2250
ATATATAAAG	ACATTGTTCT	GAGAGCAGTA	CGATTTTCATG	GAAAGATTGT	2300
TTGGTGGCTT	TGTTAAAATT	AATAAAGAAT	TTTTAAGGAT	ATAGTGTAAT	2350
TTTCTTCATT	GCATTAATAT	AACCAAATAT	GCCTACCTAT	CTTTGTCTTG	2400
AACCAAATGA	ATAGATTTGG	AATACTTTAT	TGTAATTGAA	TTTGATATAA	2450

```

AGTTGACTGA GCATTTATGT GTTACCTGCA TGCTTCTGGG TGCATTGAAA 2500
TATTTTAACT TTTAAATGA TACTATGTTG TTTCAATTTT GACTACCTTT 2550
TGTGAGGCAT ACTGGCTACC TCCTCCTATT AGCTAAGATC TTCCAAAGCC 2600
TTATAATGAA AAGTTTATAT AAACCATTTC TCTTTCAAAT CACTGTCATA 2650
CTTGGTCACG GATCCCAGGA ATATTGTAAA TTTTCTAATT TACTCTGCAC 2700
TTTGTATATC CAGCCTCTAT TACCCTCAAG GTGAATATAA AACTATGTCT 2750
TTTGAATATT TCTCTTTGAT TTTGTGATAG CAGTCCCTCA TATCTTGTAC 2800
TAATTTTATG TATATGTCAA CAGTGGTTGG TCTTTAAAAA TAAATCAAAG 2850
AATAAGTAAA AAAAAAAAAA AAAAAAAAAA AAAAAATAAA AAAAAAAAAA 2900
A 2901

```

(2) INFORMATION FOR SEQUENCE ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH:      626 amino acids
(B) TYPE:        amino acid
(C) STRANDNESS:  single
(D) TOPOLOGY:    linear

```

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```

Met Val Val Gln Thr Arg Phe Pro Ser Trp Ile Ile Leu Cys Tyr Ile
1           5           10           15
Trp Leu Leu Gly Phe Ala His Thr Gly Glu Ala Gln Ala Ala Lys Glu
20          25          30
Val Leu Leu Leu Asp Ser Lys Ala Gln Gln Thr Glu Leu Glu Trp Ile
35          40          45
Ser Ser Pro Pro Ser Gly Trp Glu Glu Ile Ser Gly Leu Asp Glu Asn
50          55          60
Tyr Thr Pro Ile Arg Thr Tyr Gln Val Cys Gln Val Met Glu Pro Asn
65          70          75          80
Gln Asn Asn Trp Leu Arg Thr Asn Trp Ile Ser Lys Gly Asn Ala Gln
85          90          95
Arg Ile Phe Val Glu Leu Lys Phe Thr Leu Arg Asp Cys Asn Ser Leu
100         105         110

```

117

Pro	Gly	Val	Leu	Gly	Thr	Cys	Lys	Glu	Thr	Phe	Asn	Leu	Tyr	Tyr	Tyr	115	120	125
Glu	Thr	Asp	Tyr	Asp	Thr	Gly	Arg	Asn	Ile	Arg	Glu	Asn	Leu	Tyr	Val	130	135	140
Lys	Ile	Asp	Thr	Ile	Ala	Ala	Asp	Glu	Ser	Phe	Thr	Gln	Gly	Asp	Leu	145	150	155
Gly	Glu	Arg	Lys	Met	Lys	Leu	Asn	Thr	Glu	Val	Arg	Glu	Ile	Gly	Pro	165	170	175
Leu	Ser	Lys	Lys	Gly	Phe	Tyr	Leu	Ala	Phe	Gln	Asp	Val	Gly	Ala	Cys	180	185	190
Ile	Ala	Leu	Val	Ser	Val	Lys	Val	Tyr	Tyr	Lys	Lys	Cys	Trp	Thr	Ile	195	200	205
Val	Glu	Asn	Leu	Ala	Val	Phe	Pro	Asp	Thr	Val	Thr	Gly	Ser	Glu	Phe	210	215	220
Ser	Ser	Leu	Val	Glu	Val	Arg	Gly	Thr	Cys	Val	Ser	Ser	Ala	Glu	Glu	225	230	235
Glu	Ala	Glu	Asn	Ser	Pro	Arg	Met	His	Cys	Ser	Ala	Glu	Gly	Glu	Trp	245	250	255
Leu	Val	Pro	Ile	Gly	Lys	Cys	Ile	Cys	Lys	Ala	Gly	Tyr	Gln	Gln	Lys	260	265	270
Gly	Asp	Thr	Cys	Glu	Pro	Cys	Gly	Arg	Arg	Phe	Tyr	Lys	Ser	Ser	Ser	275	280	285
Gln	Asp	Leu	Gln	Cys	Ser	Arg	Cys	Pro	Thr	His	Ser	Phe	Ser	Asp	Arg	290	295	300
Glu	Gly	Ser	Ser	Arg	Cys	Glu	Cys	Glu	Asp	Gly	Tyr	Tyr	Arg	Ala	Pro	305	310	315
Ser	Asp	Pro	Pro	Tyr	Val	Ala	Cys	Thr	Arg	Pro	Pro	Ser	Ala	Pro	Gln	325	330	335
Asn	Leu	Ile	Phe	Asn	Ile	Asn	Gln	Thr	Thr	Val	Ser	Leu	Glu	Trp	Ser	340	345	350
Pro	Pro	Ala	Asp	Asn	Gly	Gly	Arg	Asn	Asp	Val	Thr	Tyr	Arg	Ile	Leu	355	360	365
Cys	Lys	Arg	Cys	Ser	Trp	Glu	Gln	Gly	Glu	Cys	Val	Pro	Cys	Gly	Ser	370	375	380
Asn	Ile	Gly	Tyr	Met	Pro	Gln	Gln	Thr	Gly	Leu	Glu	Asp	Asn	Tyr	Val	385	390	395

SSSD/27372. v01

118

Thr Val Met Asp Leu Leu Ala His Ala Asn Tyr Thr Phe Glu Val Glu
405 410 415
Ala Val Asn Gly Val Ser Asp Leu Ser Arg Ser Gln Arg Leu Phe Ala
420 425 430
Ala Val Ser Ile Thr Thr Gly Gln Ala Ala Pro Ser Gln Val Ser Gly
435 440 445
Val Met Lys Glu Arg Val Leu Gln Arg Ser Val Gln Leu Ser Trp Gln
450 455 460
Glu Pro Glu His Pro Asn Gly Val Ile Thr Glu Tyr Glu Ile Lys Tyr
465 470 475 480
Tyr Glu Lys Asp Gln Arg Glu Arg Thr Tyr Ser Thr Leu Lys Thr Lys
485 490 495
Ser Thr Ser Ala Ser Ile Asn Asn Leu Lys Pro Gly Thr Val Tyr Val
500 505 510
Phe Gln Ile Arg Ala Val Thr Ala Ala Gly Tyr Gly Asn Tyr Ser Pro
515 520 525
Arg Leu Asp Val Ala Thr Leu Glu Glu Ala Ser Gly Lys Met Phe Glu
530 535 540
Ala Thr Ala Val Ser Ser Glu Gln Asn Pro Val Ile Ile Ile Ala Val
545 550 555 560
Val Ala Val Ala Gly Thr Ile Ile Leu Val Phe Met Val Phe Gly Phe
565 570 575
Ile Ile Gly Arg Arg His Cys Gly Tyr Ser Lys Ala Asp Gln Glu Gly
580 585 590
Asp Glu Glu Leu Tyr Phe His Ser Leu Tyr Arg Glu Arg Gly Asp Gly
595 600 605
Met Glu Lys Thr Gln His Asn Lys Lys Trp Met Ile Ala Ser Cys Ser
610 615 620
Arg Leu
625

(2) INFORMATION FOR SEQUENCE ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2323 base pairs
(B) TYPE: nucleic acid
(C) STRANDNESS: single
(D) TOPOLOGY: linear

SSSD/27372. v01

(ii) MOLECULE TYPE: nucleic

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

AAGCGGCCGG TCTGCAGTCG GAGACTTGCA GGCAGCAAAC ACGGTGCGAA	50
CGAACCGGAG GGGGGAGAGA GAAATCAAAC AGCTAAGCGT GGAGCAGACG	100
GCCTGGGACC CAGAAGGGGA TCGATGCGAG GAGCGCAATA ATAACAACAA	150
TAATAACCCA CTTCGAGCA AACAGCATCT AAAGAGCTGC GACCCAACTG	200
CAGCCTAAAA AAATCAAACC TGCTCATGCA CCATGGTTGT TCAAACCTCGG	250
TTCCCTTCGT GGATTATTTT GTGTTACATC TGGCTGCTTG GCTTTGCACA	300
CACGGGGGAG GCGCAGGCTG CGAAGGAAGT ACTATTACTG GACTCGAAAG	350
CACAACAAAC AGAATTGGAA TGGATTTCCT CTCCACCCAG TGGGTGGGAA	400
GAAATTAGTG GTTTGGATGA GAACTACACT CCGATAAGAA CATACCAGGT	450
GTGCCAGGTC ATGGAGCCCA ACCAGAACAA CTGGCTGCGG ACTAACTGGA	500
TTTCTAAAGG CAACGCACAA AGGATTTTTG TAGAATTGAA ATTACCTTG	550
AGGGATTGTA ATAGTCTTCC CGGAGTCCTG GGAAGTTGCA AGGAAACGTT	600
TAATTTGTAC TATTATGAAA CAGACTACGA CACCGGCAGG AATATACGAG	650
AAAACCTTTA TGTTAAAATA GACACCATTG CTGCAGATGA AAGTTTCACA	700
CAAGGTGACC TTGGTGAAAG AAAGATGAAG CTGAACACTG AGGTGAGAGA	750
GATTGGACCT TTGTCCAAAA AGGGATTCTA TCTTGCCTTT CAGGATGTAG	800
GGGCTTGCAT AGCATTGGTT TCTGTCAAAG TGTACTACAA GAAGTGCTGG	850
ACCATTGTTG AGAACTTAGC TGTCTTTCCA GATACAGTGA CTGGTTCGGA	900
ATTTTCCTCC TTAGTCGAGG TCCGTGGGAC ATGTGTCAGC AGTGCCGAGG	950
AAGAGGCAGA AAATTCCCCC AGAATGCATT GCAGTGCAGA AGGAGAGTGG	1000
CTAGTACCCA TTGGAAAATG CATCTGCAAA GCAGGCTATC AGCAAAAAGG	1050
GGACACTTGC GAACCCTGTG GCCGCAGGTT CTACAAATCT TCCTCTCAGG	1100
ATCTCCAGTG TTCTCGTTGT CCAACCCACA GCTTCTCTGA CCGAGAAGGA	1150
TCATCCAGGT GTGAATGTGA AGATGGGTAC TACAGAGCTC CTTCTGATCC	1200
ACCATACGTT GCATGCACGA GGCCTCCCTC TGCACCACAG AACCTTATTT	1250
TCAATATCAA TCAAACGACT GTAAGTTTGG AATGGAGTCC TCCGGCTGAC	1300

SSSD/27372. v01

```

AACGGGGGAA GAAACGATGT CACCTACAGA ATACTGTGTA AGCGGTGCAG 1350
TTGGGAACAG GGAGAATGTG TGCCATGCGG AAGTAACATT GGATACATGC 1400
CCCAGCAGAC GGGATTAGAG GATAACTATG TCACTGTCAT GGACCTACTT 1450
GCCCATGCAA ATTACACTTT CGAAGTTGAA GCTGTAAATG GAGTTTCGGA 1500
CTTAAGCAGA TCCCAGAGGC TCTTCGCTGC TGTTAGCATC ACCACCGGTC 1550
AAGCAGCTCC CTCGCAAGTG AGTGGAGTCA TGAAGGAGCG AGTACTGCAG 1600
CGGAGTGTGC AGCTTTCCTG GCAGGAGCCG GAGCATCCCA ATGGAGTCAT 1650
CACGGAATAT GAAATCAAGT ATTATGAGAA AGATCAACGG GAAAGGACGT 1700
ACTCAACACT CAAAACCAAG TCCACCTCCG CCTCCATTAA TAATCTGAAA 1750
CCGGGAACAG TGTACGTCTT TCAGATCCGG GCGGTCACTG CTGCCGGTTA 1800
TGGAAACTAC AGCCCTAGGC TTGATGTTGC CACACTTGAG GAAGCTTCAG 1850
GTAAATGTT TGAAGCGACA GCAGTCTCCA GTGAACAGAA TCCTGTCATC 1900
ATAATTGCTG TAGTGGCTGT AGCAGGGACC ATCATCTTGG TGTTCATGGT 1950
GTTCCGGCTTC ATCATTTGAA GAAGGCACTG TGGTTATAGC AAGGCTGACC 2000
AAGAAGGGGA TGAAGAACTC TACTTTCATT CTCTTTACAG GGAAAGGGGA 2050
GACGGGATGG AAAAGACACA GCACAATAAG AAGTGGATGA TTGCATCGTG 2100
CTCTCGTTTG TAGGTCTCTT TTCCTAATCA ACACTATGAT TTTGAAGTAC 2150
GCGTACACGA AGCAAACGGG AAGAGATAAG GAATTAGCAT TGTGAACCTG 2200
ACTGTAATCC TCTCTTCCGG AAAGAGATGA GATGCTATTG CGATGAGAAT 2250
GTACAACTTG CACCTTGAAA TCTTTTTTGA TAATTAGTGC TCAGGGGAGG 2300
GGGGGGGAAG TAGAGAAAGC AAA 2323

```

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH:      6 amino acids
(B) TYPE:        amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY:    linear

```

(ii) MOLECULE TYPE: peptide

121

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Ala Ala Thr Ala Ala Ala
5

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Ala Ala Thr Ala Ala Ala
5

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

His Arg Asp Leu Ala Ala
5

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

SSSD/27372. v01

122

(ix) FEATURE:

(D) OTHER INFORMATION: Xaa in position 2 is valine
or methionine; Xaa in
position 5 is phenylalanine
or tyrosine.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Asp Xaa Trp Ser Xaa Gly
5

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 993 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Met Val Val Gln Thr Arg Phe Pro Ser Trp Ile Ile Leu Cys Tyr Ile
1 5 10 15
Trp Leu Leu Gly Phe Ala His Thr Gly Glu Ala Gln Ala Ala Lys Glu
20 25 30
Val Leu Leu Leu Asp Ser Lys Ala Gln Gln Thr Glu Leu Glu Trp Ile
35 40 45
Ser Ser Pro Pro Ser Gly Trp Glu Glu Ile Ser Gly Leu Asp Glu Asn
50 55 60
Tyr Thr Pro Ile Arg Thr Tyr Gln Val Cys Gln Val Met Glu Pro Asn
65 70 75 80
Gln Asn Asn Trp Leu Arg Thr Asn Trp Ile Ser Lys Gly Asn Ala Gln
85 90 95
Arg Ile Phe Val Glu Leu Lys Phe Thr Leu Arg Asp Cys Asn Ser Leu
100 105 110
Pro Gly Val Leu Gly Thr Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr
115 120 125
Glu Thr Asp Tyr Asp Thr Gly Arg Asn Ile Arg Glu Asn Leu Tyr Val
130 135 140

123

Lys Ile Asp Thr Ile Ala Ala Asp Glu Ser Phe Thr Gln Gly Asp Leu
 145 150 155 160
 Gly Glu Arg Lys Met Lys Leu Asn Thr Glu Val Arg Glu Ile Gly Pro
 165 170 175
 Leu Ser Lys Lys Gly Phe Tyr Leu Ala Phe Gln Asp Val Gly Ala Cys
 180 185 190
 Ile Ala Leu Val Ser Val Lys Val Tyr Tyr Lys Lys Cys Trp Thr Ile
 195 200 205
 Val Glu Asn Leu Ala Val Phe Pro Asp Thr Val Thr Gly Ser Glu Phe
 210 215 220
 Ser Ser Leu Val Glu Val Arg Gly Thr Cys Val Ser Ser Ala Glu Glu
 225 230 235 240
 Glu Ala Glu Asn Ser Pro Arg Met His Cys Ser Ala Glu Gly Glu Trp
 245 250 255
 Leu Val Pro Ile Gly Lys Cys Ile Cys Lys Ala Gly Tyr Gln Gln Lys
 260 265 270
 Gly Asp Thr Cys Glu Pro Cys Gly Arg Arg Phe Tyr Lys Ser Ser Ser
 275 280 285
 Gln Asp Leu Gln Cys Ser Arg Cys Pro Thr His Ser Phe Ser Asp Arg
 290 295 300
 Glu Gly Ser Ser Arg Cys Glu Cys Glu Asp Gly Tyr Tyr Arg Ala Pro
 305 310 315 320
 Ser Asp Pro Pro Tyr Val Ala Cys Thr Arg Pro Pro Ser Ala Pro Gln
 325 330 335
 Asn Leu Ile Phe Asn Ile Asn Gln Thr Thr Val Ser Leu Glu Trp Ser
 340 345 350
 Pro Pro Ala Asp Asn Gly Gly Arg Asn Asp Val Thr Tyr Arg Ile Leu
 355 360 365
 Cys Lys Arg Cys Ser Trp Glu Gln Gly Glu Cys Val Pro Cys Gly Ser
 370 375 380
 Asn Ile Gly Tyr Met Pro Gln Gln Thr Gly Leu Glu Asp Asn Tyr Val
 385 390 395 400
 Thr Val Met Asp Leu Leu Ala His Ala Asn Tyr Thr Phe Glu Val Glu
 405 410 415
 Ala Val Asn Gly Val Ser Asp Leu Ser Arg Ser Gln Arg Leu Phe Ala
 420 425 430

124

Ala Val Ser Ile Thr Thr Gly Gln Ala Ala Pro Ser Gln Val Ser Gly
435 440 445

Val Met Lys Glu Arg Val Leu Gln Arg Ser Val Gln Leu Ser Trp Gln
450 455 460

Glu Pro Glu His Pro Asn Gly Val Ile Thr Glu Tyr Glu Ile Lys Tyr
465 470 475 480

Tyr Glu Lys Asp Gln Arg Glu Arg Thr Tyr Ser Thr Leu Lys Thr Lys
485 490 495

Ser Thr Ser Ala Ser Ile Asn Asn Leu Lys Pro Gly Thr Val Tyr Val
500 505 510

Phe Gln Ile Arg Ala Val Thr Ala Ala Gly Tyr Gly Asn Tyr Ser Pro
515 520 525

Arg Leu Asp Val Ala Thr Leu Glu Glu Ala Ser Ala Thr Ala Val Ser
530 535 540

Ser Glu Gln Asn Pro Val Ile Ile Ile Ala Val Val Ala Val Ala Gly
545 550 555 560

Thr Ile Ile Leu Val Phe Met Val Phe Gly Phe Ile Ile Gly Arg Arg
565 570 575

His Cys Gly Tyr Ser Lys Ala Asp Gln Glu Gly Asp Glu Glu Leu Tyr
580 585 590

Phe His Phe Lys Phe Pro Gly Thr Lys Thr Tyr Ile Asp Pro Glu Thr
595 600 605

Tyr Glu Asp Pro Asn Arg Ala Val His Gln Phe Ala Lys Glu Leu Asp
610 615 620

Ala Ser Cys Ile Lys Ile Glu Arg Val Ile Gly Ala Gly Glu Phe Gly
625 630 635 640

Glu Val Cys Ser Gly Arg Leu Lys Leu Pro Gly Gln Arg Asp Val Ala
645 650 655

Val Ala Ile Lys Thr Leu Lys Val Gly Tyr Thr Glu Lys Gln Arg Arg
660 665 670

Asp Phe Leu Cys Glu Ala Ser Ile Met Gly Gln Phe Asp His Pro Asn
675 680 685

Val Val His Leu Glu Gly Val Val Thr Arg Gly Lys Pro Val Met Ile
690 695 700

Val Ile Glu Phe Met Glu Asn Gly Ala Leu Asp Ala Phe Leu Arg Lys
705 710 715 720

SSSD/27372. v01

222803001 01 19061 552 0061 10 17033084242

NOV 24 1956 10:04 FR LYON*KLION

125

His Asp Gly Gln Phe Thr Val Ile Gln Leu Val Gly Met Leu Arg Gly
725 730 735
Ile Ala Ala Gly Met Arg Tyr Leu Ala Asp Met Gly Tyr Val His Arg
740 745 750
Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser Asn Leu Val Cys Lys
755 760 765
Val Ser Asp Phe Gly Leu Ser Arg Val Ile Glu Asp Asp Pro Glu Ala
770 775 780
Val Tyr Thr Thr Thr Gly Gly Lys Ile Pro Val Arg Trp Thr Ala Pro
785 790 795 800
Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val Trp Ser
805 810 815
Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu Arg Pro Tyr
820 825 830
Trp Asp Met Ser Asn Gln Asp Val Ile Lys Ala Ile Glu Glu Gly Tyr
835 840 845
Arg Leu Pro Ala Pro Met Asp Cys Pro Ala Gly Leu His Gln Leu Met
850 855 860
Leu Asp Cys Trp Gln Lys Asp Arg Ala Glu Arg Pro Lys Phe Glu Gln
865 870 875 880
Ile Val Gly Ile Leu Asp Lys Met Ile Arg Asn Pro Ser Ser Leu Lys
885 890 895
Thr Pro Leu Gly Thr Cys Ser Arg Pro Leu Ser Pro Leu Leu Asp Gln
900 905 910
Ser Thr Pro Asp Phe Thr Ala Phe Cys Ser Val Gly Glu Trp Leu Gln
915 920 925
Ala Ile Lys Met Glu Arg Tyr Lys Asp Asn Phe Thr Ala Ala Gly Tyr
930 935 940
Asn Ser Leu Glu Ser Val Ala Arg Met Thr Ile Asp Asp Val Met Ser
945 950 955 960
Leu Gly Ile Thr Leu Val Gly His Gln Lys Lys Ile Met Ser Ser Ile
965 970 975
Gln Thr Met Arg Ala Gln Met Leu His Leu His Gly Thr Gly Ile Gln
980 985 990
Val

SSSD/27372. v01

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 994 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

```

Met Val Val Gln Thr Arg Phe Pro Ser Trp Ile Ile Leu Cys Tyr Ile
 1             5             10             15
Trp Leu Leu Gly Phe Ala His Thr Gly Glu Ala Gln Ala Ala Lys Glu
      20             25             30
Val Leu Leu Leu Asp Ser Lys Ala Gln Gln Thr Glu Leu Glu Trp Ile
      35             40             45
Ser Ser Pro Pro Ser Gly Trp Glu Glu Ile Ser Gly Leu Asp Glu Asn
      50             55             60
Tyr Thr Pro Ile Arg Thr Tyr Gln Val Cys Gln Val Met Glu Pro Asn
      65             70             75             80
Gln Asn Asn Trp Leu Arg Thr Asn Trp Ile Ser Lys Gly Asn Ala Gln
      85             90             95
Arg Ile Phe Val Glu Leu Lys Phe Thr Leu Arg Asp Cys Asn Ser Leu
      100            105            110
Pro Gly Val Leu Gly Thr Cys Lys Glu Thr Phe Asn Leu Tyr Tyr Tyr
      115            120            125
Glu Thr Asp Tyr Asp Thr Gly Arg Asn Ile Arg Glu Asn Leu Tyr Val
      130            135            140
Lys Ile Asp Thr Ile Ala Ala Asp Glu Ser Phe Thr Gln Gly Asp Leu
      145            150            155            160
Gly Glu Arg Lys Met Lys Leu Asn Thr Glu Val Arg Glu Ile Gly Pro
      165            170            175
Leu Ser Lys Lys Gly Phe Tyr Leu Ala Phe Gln Asp Val Gly Ala Cys
      180            185            190
Ile Ala Leu Val Ser Val Lys Val Tyr Tyr Lys Lys Cys Trp Thr Ile
      195            200            205

```


Val Glu Asn Leu Ala Val Phe Pro Asp Thr Val Thr Gly Ser Glu Phe
 210 215 220
 Ser Ser Leu Val Glu Val Arg Gly Thr Cys Val Ser Ser Ala Glu Glu
 225 230 235 240
 Glu Ala Glu Asn Ser Pro Arg Met His Cys Ser Ala Glu Gly Glu Trp
 245 250 255
 Leu Val Pro Ile Gly Lys Cys Ile Cys Lys Ala Gly Tyr Gln Gln Lys
 260 265 270
 Gly Asp Thr Cys Glu Pro Cys Gly Arg Arg Phe Tyr Lys Ser Ser Ser
 275 280 285
 Gln Asp Leu Gln Cys Ser Arg Cys Pro Thr His Ser Phe Ser Asp Arg
 290 295 300
 Glu Gly Ser Ser Arg Cys Glu Cys Glu Asp Gly Tyr Tyr Arg Ala Pro
 305 310 315 320
 Ser Asp Pro Pro Tyr Val Ala Cys Thr Arg Pro Pro Ser Ala Pro Gln
 325 330 335
 Asn Leu Ile Phe Asn Ile Asn Gln Thr Thr Val Ser Leu Glu Trp Ser
 340 345 350
 Pro Pro Ala Asp Asn Gly Gly Arg Asn Asp Val Thr Tyr Arg Ile Leu
 355 360 365
 Cys Lys Arg Cys Ser Trp Glu Gln Gly Glu Cys Val Pro Cys Gly Ser
 370 375 380
 Asn Ile Gly Tyr Met Pro Gln Gln Thr Gly Leu Glu Asp Asn Tyr Val
 385 390 395 400
 Thr Val Met Asp Leu Leu Ala His Ala Asn Tyr Thr Phe Glu Val Glu
 405 410 415
 Ala Val Asn Gly Val Ser Asp Leu Ser Arg Ser Gln Arg Leu Phe Ala
 420 425 430
 Ala Val Ser Ile Thr Thr Gly Gln Ala Ala Pro Ser Gln Val Ser Gly
 435 440 445
 Val Met Lys Glu Arg Val Leu Gln Arg Ser Val Gln Leu Ser Trp Gln
 450 455 460
 Glu Pro Glu His Pro Asn Gly Val Ile Thr Glu Tyr Glu Ile Lys Tyr
 465 470 475 480
 Tyr Glu Lys Asp Gln Arg Glu Arg Thr Tyr Ser Thr Leu Lys Thr Lys
 485 490 495

128

Ser Thr Ser Ala Ser Ile Asn Asn Leu Lys Pro Gly Thr Val Tyr Val
500 505 510

Phe Gln Ile Arg Ala Val Thr Ala Ala Gly Tyr Gly Asn Tyr Ser Pro
515 520 525

Arg Leu Asp Val Ala Thr Leu Glu Glu Ala Ser Gly Lys Met Phe Glu
530 535 540

Ala Thr Ala Val Ser Ser Glu Gln Asn Pro Val Ile Ile Ile Ala Val
545 550 555 560

Val Ala Val Ala Gly Thr Ile Ile Leu Val Phe Met Val Phe Gly Phe
565 570 575

Ile Ile Gly Arg Arg His Cys Gly Tyr Ser Lys Ala Asp Gln Glu Gly
580 585 590

Asp Glu Glu Leu Tyr Phe His Cys Thr Lys Thr Tyr Ile Asp Pro Glu
595 600 605

Thr Tyr Glu Asp Pro Asn Arg Ala Val His Gln Phe Ala Lys Glu Leu
610 615 620

Asp Ala Ser Cys Ile Lys Ile Glu Arg Val Ile Gly Ala Gly Glu Phe
625 630 635 640

Gly Glu Val Cys Ser Gly Arg Leu Lys Leu Pro Gly Gln Arg Asp Val
645 650 655

Ala Val Ala Ile Lys Thr Leu Lys Val Gly Tyr Thr Glu Lys Gln Arg
660 665 670

Arg Asp Phe Leu Cys Glu Ala Ser Ile Met Gly Gln Phe Asp His Pro
675 680 685

Asn Val Val His Leu Glu Gly Val Val Thr Arg Gly Lys Pro Val Met
690 695 700

Ile Val Ile Glu Phe Met Glu Asn Gly Ala Leu Asp Ala Phe Leu Arg
705 710 715 720

Lys His Asp Gly Gln Phe Thr Val Ile Gln Leu Val Gly Met Leu Arg
725 730 735

Gly Ile Ala Ala Gly Met Arg Tyr Leu Ala Asp Met Gly Tyr Val His
740 745 750

Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Asn Ser Asn Leu Val Cys
755 760 765

Lys Val Ser Asp Phe Gly Leu Ser Arg Val Ile Glu Asp Asp Pro Glu
770 775 780

SSSD/27372. v01

** TOTAL PAGE.31 **

129

Ala Val Tyr Thr Thr Thr Gly Gly Lys Ile Pro Val Arg Trp Thr Ala
785 790 795 800

Pro Glu Ala Ile Gln Tyr Arg Lys Phe Thr Ser Ala Ser Asp Val Trp
805 810 815

Ser Tyr Gly Ile Val Met Trp Glu Val Met Ser Tyr Gly Glu Arg Pro
820 825 830

Tyr Trp Asp Met Ser Asn Gln Asp Val Ile Lys Ala Ile Glu Glu Gly
835 840 845

Tyr Arg Leu Pro Ala Pro Met Asp Cys Pro Ala Gly Leu His Gln Leu
850 855 860

Met Leu Asp Cys Trp Gln Lys Asp Arg Ala Glu Arg Pro Lys Phe Glu
865 870 875 880

Gln Ile Val Gly Ile Leu Asp Lys Met Ile Arg Asn Pro Ser Ser Leu
885 890 895

Lys Thr Pro Leu Gly Thr Cys Ser Arg Pro Leu Ser Pro Leu Leu Asp
900 905 910

Gln Ser Thr Pro Asp Phe Thr Ala Phe Cys Ser Val Gly Glu Trp Leu
915 920 925

Gln Ala Ile Lys Met Glu Arg Tyr Lys Asp Asn Phe Thr Ala Ala Gly
930 935 940

Tyr Asn Ser Leu Glu Ser Val Ala Arg Met Thr Ile Asp Asp Val Met
945 950 955 960

Ser Leu Gly Ile Thr Leu Val Gly His Gln Lys Lys Ile Met Ser Ser
965 970 975

Ile Gln Thr Met Arg Ala Gln Met Leu His Leu His Gly Thr Gly Ile
980 985 990

Gln Val